On algebraic time-derivative estimation and deadbeat state reconstruction

Johann Reger[†]

Abstract

This note places into perspective the so-called algebraic time-derivative estimation method recently introduced by Fliess and co-authors with standard results from linear state-space theory for control systems. In particular, it is shown that the algebraic method can in a sense be seen as a special case of deadbeat state estimation based on the reconstructibility Gramian of the considered system.

I. Introduction

In the past few years, the algebraic approach to estimation in control systems proposed by Fliess and co-workers has generated a number of interesting results for different problems of estimation of dynamical systems such as state estimation, parametric identification, and fault diagnosis, to name but a few (see [10], [8], [6], [5] and references therein). Loosely speaking, this new estimation approach is mainly based on the robust computation of the time-derivatives of a noisy signal by using a finite weighted combination of time-integrations of this signal. These results, obtained through the use of differential algebra and operational calculus [17], allow to obtain an estimate of the time-derivative of a particular order in an arbitrary small amount of time [9].

Questions arise on how to relate the above to more classical results of automatic control, and in particular to linear system theory. The present paper contributes to this discussion by showing that the algebraic time-derivative estimation method, as presented in [18] and references therein, can be seen, in a sense, as a special case of previously known state-space results exhibiting a deadbeat property.

[†]Johann Reger is a professor and head of the Control Engineering Group, Computer Science and Automation Faculty, Ilmenau University of Technology, P.O. Box 10 05 65, D-98684 Ilmenau, Germany (e-mail: reger@ieee.org).

After this introduction, we briefly recall in Section II the main results of the algebraic time-derivative estimation method. Then, in Section III, we recall a few results of linear observability theory and show how in particular the reconstructibility Gramian can be related to the algebraic method. We end this paper with a few additional remarks on how to relate further extensions of the algebraic approach with different areas of control systems theory.

II. Algebraic time-derivative estimation

The algebraic derivative estimation techniques have been presented in various styles and frameworks, mostly based on abstract algebra and operational calculus. Because of its practical interest, we recall here only the main result for a moving-horizon version of the approach (see [18] and [31]). However, note that the results shown in the present paper would also be very easily applicable to earlier expanding-horizon versions that can be found in [4] or [10].

Consider a real-valued, N-th degree polynomial function of time

$$y(t) = \sum_{i=0}^{N} \frac{a_i}{i!}\, t^i \tag{1}$$

where the terms a_i are unknown constant coefficients. The goal is to obtain estimates of the time-derivatives of $y(t)$, up to order N.

In [4], [3], [19], Fliess and co-workers proposed to do so by, roughly speaking, resorting to algebraic combinations of moving-horizon time-integrations of the available signal $y(t)$. Let us briefly recall these results in the following theorem [18], [31].

Theorem 1: For all $t \geq T$, the j-th order time-derivative estimate $\hat{y}^{(j)}(t)$, $j = 0, 1, 2, \ldots, N$, of the polynomial signal $y(t)$ as defined in (1) satisfies the convolution

$$\hat{y}^{(j)}(t) = \int_0^T H_j(T,\tau)\, y(t-\tau)\, d\tau\,, \quad j = 0, 1, \ldots, N \tag{2}$$

where the convolution kernel

$$H_j(T,\tau) = \frac{(N+j+1)!\,(N+1)!}{T^{N+j+1}} \sum_{\kappa_1=0}^{N-j} \sum_{\kappa_2=0}^{j} \frac{(T-\tau)^{\kappa_1+\kappa_2}\,(-\tau)^{N-\kappa_1-\kappa_2}}{\kappa_1!\kappa_2!(N-j-\kappa_1)!(j-\kappa_2)!(N-\kappa_1-\kappa_2)!(\kappa_1+\kappa_2)!(N-\kappa_1+1)} \tag{3}$$

depends on the order j of the time derivative to be estimated and on an arbitrary constant time window length $T > 0$. $\square$

For the interested reader, as well as for the sake of completeness, a way to derive the results of Theorem 1 is given in Appendix A.

Thus, considering for example the degree-one polynomial

$$y(t) = a_0 + a_1\, t \tag{4}$$

applying Theorem 1 would simply give us the following first-order time-derivative estimate

$$\hat{\dot{y}}(t) = \int_0^T \frac{6}{T^3}\big(T - 2\tau\big)\, y(t-\tau)\, d\tau\,. \tag{5}$$

The effect of the time-integration is obviously to dampen the impact of the measurement noise on the estimate. Note that this feature can also be used to filter out noise from the signal $y(t)$ itself, as the zero-order time-derivative estimator would be

$$\hat{y}(t) = \int_0^T \frac{2}{T^2}\big(2T - 3\tau\big)\, y(t-\tau)\, d\tau \tag{6}$$

as obtained, once again, from Theorem 1.

III. From deadbeat reconstruction of the state to the algebraic method

As will be seen, the above may be related in several ways to more traditional results of classical linear control theory. To do so, consider now the following linear time-varying system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) \tag{7}$$

$$y(t) = \mathbf{C}(t)\mathbf{x}(t) \tag{8}$$

where $\mathbf{x}(t) \in \mathbb{R}^{N+1}$ and $y(t) \in \mathbb{R}$. Note that while the form of system (7)-(8) was chosen for the sake of simplicity and ease of presentation, the discussion of the present section is extendible to systems with multiple inputs and outputs.

Then let us briefly recall a few elements pertaining to the notion of state *reconstructibility* [11], [1], [20]. As noted in Willems and Mitter [30], this property has been quite overlooked in the control literature, possibly because of its equivalence with observability for linear continuous-time systems. Loosely speaking, we say that system (7)-(8) is *reconstructible* on $[t_0, t_1]$ if $\mathbf{x}(t_1)$ can be obtained from the measurements $y(t)$ for $t \in [t_0, t_1]$.

A standard way of determining $\mathbf{x}(t_1)$ can be obtained by first writing the following expression for the output

$$y(\tau) = \mathbf{C}(\tau)\, \mathbf{\Phi}(\tau, t_1)\, \mathbf{x}(t_1) \tag{9}$$

where $\mathbf{\Phi}(\tau, t)$ is the transition matrix of (7). Then, left-multiply and integrate (9) to get

$$\int_{t_0}^{t_1} \mathbf{\Phi}^{\mathrm{T}}(\tau, t_1)\, \mathbf{C}^{\mathrm{T}}(\tau)\, y(\tau)\, d\tau = \left(\int_{t_0}^{t_1} \mathbf{\Phi}^{\mathrm{T}}(\tau, t_1)\, \mathbf{C}^{\mathrm{T}}(\tau)\, \mathbf{C}(\tau)\, \mathbf{\Phi}(\tau, t_1)\, d\tau \right) \mathbf{x}(t_1) \tag{10}$$

Since in eq. (10) $\mathbf{x}(t_1)$ is a constant term with respect to the integral, it can be isolated, and we finally get, for an estimate $\hat{\mathbf{x}}(t_1)$ of $\mathbf{x}(t_1)$,

$$\hat{\mathbf{x}}(t_1) := \mathbf{W}_{\mathrm{r}}^{-1}(t_0, t_1) \int_{t_0}^{t_1} \mathbf{\Phi}^{\mathrm{T}}(\tau, t_1)\, \mathbf{C}^{\mathrm{T}}(\tau)\, y(\tau)\, d\tau \tag{11}$$

where

$$\mathbf{W}_{\mathrm{r}}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{\Phi}^{\mathrm{T}}(\tau, t_1)\, \mathbf{C}^{\mathrm{T}}(\tau)\, \mathbf{C}(\tau)\, \mathbf{\Phi}(\tau, t_1)\, d\tau \tag{12}$$

is the *reconstructibility Gramian*.

In treatments of observability in textbooks, developments such as the above are mostly used, through the observability counterpart of (12), to check whether a system is observable (resp. reconstructible) or not. However, as noted in [2, p. 158] for the observability case, expression (12) can also be used to actually compute $\hat{\mathbf{x}}(t_1)$ as integration will smooth out high-frequency noise.

The above results are well-known, even if not as much used for state estimation as linear asymptotic observers are. But the former has the interesting property of allowing to give an estimate of $\mathbf{x}(t_1)$ in a *finite* time, whose value is decided by the invertibility of (12).

Interestingly, these two features of the above Gramian-based estimation – deadbeat property and time-integration, coincide with those of algebraic time-derivative estimation.

Let us push the comparison a little further in a simple way by first noticing that the degree-one polynomial (4) can be put into state-space phase-variable form with matrices

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \tag{13}$$

state $\mathbf{x}(t) = (y(t), \dot{y}(t))^{\mathrm{T}}$ and initial conditions $\mathbf{x}(0) = (a_0, a_1)^{\mathrm{T}}$.

Then, compute an estimate of $\mathbf{x}(t)$ using (11) and (12). To do so, use the fact that the matrices in (13) are time-invariant and that $\mathbf{A}^2 = \mathbf{0}$ to obtain

$$\mathbf{\Phi}(\tau, t_1) = e^{\mathbf{A}(\tau - t_1)} = \mathbf{I} + (\tau - t_1)\, \mathbf{A} = \begin{pmatrix} 1 & \tau - t_1 \\ 0 & 1 \end{pmatrix} \tag{14}$$

which implies that

$$\mathbf{C}\,\boldsymbol{\Phi}(\tau,t_1)=\begin{pmatrix}1 & \tau-t_1\end{pmatrix}\,. \tag{15}$$

Letting $t_0=t-T$ (with $T>0$ fixed) and $t_1=t$, we then obtain from (12) the following Gramian

$$\mathbf{W}_{\mathrm{r}}(t-T,t)=\begin{pmatrix}T & -\frac{T^2}{2}\\ -\frac{T^2}{2} & \frac{T^3}{3}\end{pmatrix} \tag{16}$$

which in turn is used, in combination with (11), to get

$$\hat{\mathbf{x}}(t)=\begin{pmatrix}\hat{y}(t)\\ \hat{\dot{y}}(t)\end{pmatrix}=\begin{pmatrix}\frac{4}{T} & \frac{6}{T^2}\\ \frac{6}{T^2} & \frac{12}{T^3}\end{pmatrix}\int_{t-T}^{t}\begin{pmatrix}1\\ \tau-t\end{pmatrix}y(\tau)\,d\tau\,. \tag{17}$$

Hence, similarly to the previous section, an estimate of the derivatives of a degree-one polynomial can be obtained with time-integrations of the measured signal, albeit this time using tools from classical control theory.

Note, interestingly, that in this particular example, there is more than a mere similarity. Indeed, after a simple change of variable $\sigma=t-\tau$ in (17), we find exactly the same expressions as (5) and (6).

The above second-order case can be generalized to obtain the j-th time-derivative of any polynomial simply by specializing $\mathbf{A}(t)$ and $\mathbf{C}(t)$ in (7)-(8) to get a state-space description of polynomial (1), which yields, in phase-variable form the $N+1$ square matrix

$$\mathbf{A}=\begin{pmatrix}0 & 1 & 0 & 0 & \cdots & 0\\ 0 & 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & 0\\ 0 & 0 & 0 & 0 & \cdots & 1\\ 0 & 0 & 0 & 0 & \cdots & 0\end{pmatrix} \tag{18}$$

and the $N+1$ row vector

$$\mathbf{C}=\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix} \tag{19}$$

associated to the state vector $\mathbf{x}(t)=(y(t),\dot{y}(t),...,y^{(j)}(t),...,\,y^{(N)}(t))^{\mathrm{T}}$.

After several steps in line with the previous second-order example, we obtain, similarly to Section II, an expression of the j-th time-derivative of a polynomial signal (1) based on the reconstructibility Gramian. This is summarized in the following theorem.

Theorem 2: For all $t \geq T$, the j-th order time-derivative estimate $\hat{y}^{(j)}(t)$, $j = 0, 1, 2, \ldots, N$, of the polynomial signal $y(t)$ as defined in (1) satisfies the convolution

$$\hat{y}^{(j)}(t) = \int_0^T G_j(T, \sigma)\, y(t - \sigma)\, d\sigma\,, \quad j = 0, 1, \ldots, N \tag{20}$$

where the convolution kernel

$$G_j(T, \tau) = \frac{(N+j+1)!}{T^{j+1} j!(N-j)!} \sum_{k=0}^{N} \frac{(-1)^k (N+k+1)!}{(j+k+1)(N-k)!(k!)^2} \left(\frac{\sigma}{T}\right)^k \tag{21}$$

depends on the order j of the time derivative to be estimated and on an arbitrary constant time window length $T > 0$. $\square$

Proof: In the main, the proof is based on obtaining a closed-form expression corresponding to equations (11) and (12) for the particular case with matrices (18) and (19).

Since this system is LTI, the corresponding transition matrix results from the matrix exponential of (18), i.e.

$$e^{\mathbf{A}\, t} = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 & \cdots & t^N/N! \\ 0 & 1 & t & t^2/2 & \cdots & t^{N-1}/(N-1)! \\ 0 & 0 & 1 & t & \cdots & t^{N-2}/(N-2)! \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{22}$$

which is then used to obtain the state-transition matrix

$$\mathbf{\Phi}(\tau, t_1) = e^{\mathbf{A}\,(\tau - t_1)}\,. \tag{23}$$

Consequently, the entries of the $(N+1) \times (N+1)$ reconstructibility Gramian matrix (12) read

$$[W_{\mathrm{r}}]_{ij}(t_0, t_1) = \int_{t_0}^{t_1} \frac{(\tau - t_1)^{i+j-2}}{(i-1)!(j-1)!} d\tau = \frac{-(t_0 - t_1)^{i+j-1}}{(i-1)!(j-1)!(i+j-1)}\,. \tag{24}$$

In view of (11), the inversion of this Gramian is required. Its entries are provided in closed-form by Lemma 1 in Appendix B, that is

$$\left[W_{\mathrm{r}}^{-1}\right]_{ij}(t_0, t_1) = \frac{(i-1)!\,(j-1)!\,(i+j-1)}{(t_1 - t_0)^{i+j-1}} \binom{N+i}{N+1-j} \binom{N+j}{N+1-i} \binom{i+j-2}{i-1}^2. \tag{25}$$

Hence, by using eq. (25) regarding the particular form of the transition matrix (23), the $(i+1)$-th component of $\hat{\mathbf{x}}(t)$ follows from eq. (11)

$$\hat{x}_{i+1}(t_1) = \int_{t_0}^{t_1} \sum_{j=0}^{N} \left[W_{\mathrm{r}}^{-1}\right]_{i+1,j+1}(t_0, t_1)\, \frac{(\tau - t_1)^j}{j!}\, y(\tau)\, d\tau\,. \tag{26}$$

In other words, the j-th time-derivative estimate of $y(t)$ at time $t = t_1$ can be obtained from the convolution

$$y^{(j)}(t_1) = \int_{t_0}^{t_1} \bar{G}_j(t_1, t_0, \tau)\, y(\tau)\, d\tau\,, \quad j = 0, 1, \ldots, N \tag{27}$$

where

$$\bar{G}_j(t_1, t_0, \tau) = \frac{(N+j+1)!}{(t_1-t_0)^{j+1} j!(N-j)!} \sum_{k=0}^{N} \frac{(-1)^k (N+k+1)!}{(j+k+1)(N-k)!(k!)^2} \left(\frac{t_1-\tau}{t_1-t_0}\right)^k . \tag{28}$$

A receding-horizon version of equation (27) can then be obtained as follows: Let $t_0 = t - T$ (with $T > 0$ fixed), and $t_1 = t$. Proceed then to the change of variable $\sigma = t - \tau$ to obtain (20) and (21), which completes the proof of the theorem. ■

As might be expected from the above discussion and the second-order example, it is possible to show an equivalence between the algebraic estimator of Section II and the one of Theorem 2, and this for all N. We make this statement precise in the following theorem.

Theorem 3: Let $H_j(T, \tau)$ and $G_j(T, \tau)$ be defined as in (3) and (21), respectively. Then for $T > 0$, $\tau \in [0, T]$, and $N \in \{0, 1, 2, \ldots\}$

$$H_j(T, \tau) = G_j(T, \tau), \quad j \in \{0, 1, 2, \ldots, N\}\,. \tag{29}$$

□

Proof: Theorem 3 follows from Riesz' representation theorem [22], which states that for every continuous linear functional f on a Hilbert space $\mathcal{H}$, a unique $p \in \mathcal{H}$ exists such that

$$f(q) = \langle p, q \rangle \qquad \forall q \in \mathcal{H}\,, \tag{30}$$

where $\langle . \, , . \rangle$ denotes the inner product on $\mathcal{H}$.

In order to prepare the ground for applying this theorem, first note that for parameter $T > 0$ fixed, the expressions $H_j(T, \tau)$ and $G_j(T, \tau)$, given by (3) and (21), are polynomials in τ of degree N. For t fixed, furthermore $y(t - \tau)$ is a polynomial in τ of degree N which in view of (1) consequently spans $\mathcal{H}_N$, i.e. the Hilbert space of degree N polynomials equipped with the real-valued inner product

$$\langle p, q \rangle := \int_0^T p(\tau) q(\tau)\, d\tau\,, \quad p, q \in \mathcal{H}_N\,. \tag{31}$$

Hence, for $T > 0$ fixed, $H_j(T,\tau) \in \mathcal{H}_N$ and $G_j(T,\tau) \in \mathcal{H}_N$. Moreover, letting $q(\tau) := y(t-\tau)$ with fixed $t \geq T$ we have that $q \in \mathcal{H}_N$.

In accordance with (2) and (20), let

$$f_{H_j}(q) := \int_0^T H_j(T,\tau)\, q(\tau)\, d\tau \tag{32}$$

and

$$f_{G_j}(q) := \int_0^T G_j(T,\tau)\, q(\tau)\, d\tau \tag{33}$$

for $j = 0, 1, 2, \ldots, N$.

Consequently, Theorems 1 and 2 imply that for any $q \in \mathcal{H}_N$

$$f_{H_j}(q) = f_{G_j}(q)\,, \quad j = 0, 1, 2, \ldots, N\,. \tag{34}$$

Since $H_j(T,\tau) \in \mathcal{H}_N$ and $G_j(T,\tau) \in \mathcal{H}_N$, for $T > 0$ fixed, the uniqueness of p in Riesz' theorem shows that

$$H_j(T,\tau) \equiv G_j(T,\tau) \tag{35}$$

for $j = 0, 1, 2, \ldots, N$, under the assumptions of Theorem 3. ∎

Note that other proofs of the previous theorem are also possible. For example, a somewhat more component-wise proof, based on modern computer algebra proof techniques [29], is presented in [21] by showing specifically how the terms in (3) relate to those of (21).

IV. Additional remarks

In addition to the main result of Section II, Fliess et al. proposed several extensions or modifications, several of which having also connections with different areas of control systems. Let us briefly consider some of them in the few following remarks.

For instance, note that an *expanding-horizon* version of the algebraic method was first introduced in [9], which would correspond to let $t_0 = 0$ and $t_1 = t$ in the reconstructibility Gramian perspective. In this case, an equivalence similar to Theorem 3 can still be obtained. Furthermore, note that, interestingly, letting $\mathbf{S}(t) := \mathbf{W}_{\mathrm{r}}(0,t)$, and differentiating respectively $\mathbf{S}(t)$ and the product $\mathbf{S}(t)\,\hat{\mathbf{x}}(t)$ with respect to time using a few standard manipulations, we obtain

$$\dot{\mathbf{S}}(t) = -\mathbf{A}^{\mathrm{T}}(t)\mathbf{S}(t) - \mathbf{S}(t)\mathbf{A}(t) + \mathbf{C}^{\mathrm{T}}(t)\mathbf{C}(t) \tag{36}$$

and

$$\dot{\hat{\mathbf{x}}}(t) = \left(\mathbf{A}(t) - \mathbf{S}^{-1}(t)\mathbf{C}^{\mathrm{T}}(t)\mathbf{C}(t)\right)\hat{\mathbf{x}}(t) + \mathbf{S}^{-1}(t)\mathbf{C}^{\mathrm{T}}(t)y(t) \tag{37}$$

which draw similarities with the information form of the continuous-time Kalman filter [16], [12] for system (7)-(8) with additive noise $v(t) \in \mathbb{R}$ of identity covariance, $\mathbf{R} = \mathbf{I}$, on the measurement equation (8). This in turn shows that, thanks to a simple modification of Theorem 3 for expanding horizons, links with optimal estimation could be obtained even though the derivations and motivations for the algebraic method are clearly different (see in particular [9]).

As another example, one could also consider identification problems, and parallels to the work of Fliess and Sira-Ramírez [8] on identifying linear systems using the algebraic method. Indeed, replacing (7)-(8) with

$$\dot{\boldsymbol{\varphi}}(t) = \mathbf{0} \tag{38}$$

$$y(t) = \boldsymbol{\omega}^{\mathrm{T}}(t)\,\boldsymbol{\varphi} \tag{39}$$

where $\boldsymbol{\varphi}$ is a constant parameter vector to be estimated, introduce, similar to Section III, the Gramian-based receding-horizon estimator

$$\hat{\boldsymbol{\varphi}}(t) = \mathbf{W}_{\mathrm{r}}^{-1}(t-T,t)\int_{t-T}^{t} \boldsymbol{\omega}(\tau)\,y(\tau)d\tau \tag{40}$$

with

$$\mathbf{W}_{\mathrm{r}}(t_0,t_1) = \int_{t_0}^{t_1} \boldsymbol{\omega}(\tau)\,\boldsymbol{\omega}^{\mathrm{T}}(\tau)d\tau \tag{41}$$

which is reminiscent of the well-known condition for persistent excitation (see for example [13], [15], [24]).

In an other extension presented in [19], the authors propose to further reduce the impact of measurement noise on the estimates by using additional integrations. This is also possible with the Gramian point-of-view as both sides of (10) can easily be time-integrated several additional times with respect to t_0, as opposed to only once to obtain $\mathbf{x}(t_1)$ – in fact, even filter operations with respect to the variable t_0 can be applied on both sides of (10), so as to generate a variety of further estimators. Once again, an equivalence between this result of the algebraic approach and a particularization of a reconstructibility perspective can be obtained. More generally, we can for example insert in (10) another kernel $\lambda(\tau,t_0)$ as follows

$$\hat{\mathbf{x}}(t_1) := \mathbf{W}_{\lambda}^{-1}(t_0,t_1)\int_{t_0}^{t_1} \lambda(\tau,t_0)\boldsymbol{\Phi}^{\mathrm{T}}(\tau,t_1)\,\mathbf{C}^{\mathrm{T}}(\tau)\,y(\tau)\,d\tau \tag{42}$$

where

$$\mathbf{W}_\lambda(t_0,t_1)=\int_{t_0}^{t_1}\lambda(\tau,t_0)\mathbf{\Phi}^{\mathrm{T}}(\tau,t_1)\,\mathbf{C}^{\mathrm{T}}(\tau)\,\mathbf{C}(\tau)\,\mathbf{\Phi}(\tau,t_1)\,d\tau, \tag{43}$$

this to obtain the desired response with respect to measurement noise.

Finally, and although it is clearly beyond the scope of the present paper, note that because of the convolution form of algebraic estimation (2), the latter can also be connected with Finite-Impulse Response (FIR) differentiators, on which numerous studies and results were published (see [14], [27] and references therein), with the minor difference that these differentiators are usually described in a discrete-time framework, although it is clear that a comparison similar to the present paper could also be carried out in discrete-time.

In particular, it might be of interest to compare the latest extension of the algebraic estimation approach, where time-delays are considered to improve the results, together with FIR differentiator designs considering the same issue that have been proposed over the past few years (see for example [28] and [23]).

Acknowledgments

The authors would like to express their gratitude to the associate editor, as well as to the anonymous reviewers, whose comments and suggestions were very helpful in improving the present paper. Johann Reger thanks Peter Caines and Jessy Grizzle for valuable discussions. The work of Johann Reger was supported by fellowships of the German Academic Exchange Service (DAAD), grant D/07/40582, and partially, by Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg, Germany.

APPENDIX

A. Proof of Theorem 1

The following proof resorts to standard techniques from operational calculus. To this end, we rephrase eq. (1) in the Laplace domain as

$$Y(s) = \sum_{i=0}^{N} \frac{y^{(i)}(0)}{s^{i+1}}\,, \tag{44}$$

where the coefficients a_i are identified with $y^{(i)}(0)$. In order to single out a particular term, $y^{(j)}(0)$, first multiply (44) by s^{N+1},

$$s^{N+1}\,Y(s) = \sum_{i=0}^{N} y^{(i)}(0)\,s^{N-i}\,, \tag{45}$$

which results in a polynomial form in s on the right side of (45). To eliminate the terms $y^{(j+1)}(0), \dots, y^{(N)}(0)$, differentiate (45) $N-j$ times with respect to s (see [7] for a first presentation of the idea). This yields

$$\frac{d^{N-j}}{ds^{N-j}}\left(s^{N+1}Y(s)\right) = \sum_{i=0}^{j} y^{(i)}(0)\,\frac{(N-i)!}{(j-i)!}\,s^{j-i}\,. \tag{46}$$

In the next step, we proceed to a similar treatment to eliminate the remaining constant terms $y^{(0)}(0)$, $y^{(1)}(0)$, $\dots$, $y^{(j-1)}(0)$. But before doing so, premultiply (46) by $1/s$, that is

$$\frac{1}{s}\frac{d^{N-j}}{ds^{N-j}}\left(s^{N+1}Y(s)\right) = \frac{(N-j)!}{s}y^{(j)}(0) + \sum_{i=0}^{j-1} y^{(i)}(0)\frac{(N-i)!}{(j-i)!}s^{j-i-1} \tag{47}$$

which is done to prevent $y^{(j)}(0)$ from cancelation due to a j-fold differentiation with respect to s. Indeed, the latter operation finally gives

$$\frac{d^j}{ds^j}\left(\frac{1}{s}\frac{d^{N-j}}{ds^{N-j}}\left(s^{N+1}Y(s)\right)\right) = \frac{(-1)^j\,j!\,(N-j)!}{s^{j+1}}y^{(j)}(0). \tag{48}$$

This equation could readily be transformed back into the time domain. However, the left side of (48) contains the monomial s^N, i.e. an N-fold differentiation with respect to time in the time domain, meaning if a high-frequency noise is corrupting $y(t)$, the former would be amplified as a result. Note that a similar idea can also be found in [26, p.17–18]. In order to avoid the explicit use of these time derivatives, premultiply (48) with $1/s^{N+1}$, thus implying that $y(t)$ will be integrated at least one time. Therefore, we obtain

$$\frac{1}{s^{N+1}}\frac{d^j}{ds^j}\left(\frac{1}{s}\frac{d^{N-j}}{ds^{N-j}}\left(s^{N+1}Y(s)\right)\right) = \frac{(-1)^j j!\,(N-j)!}{s^{N+j+2}}y^{(j)}(0) \tag{49}$$

where it can been seen that the term $y^{(j)}(0)$ depends only on a finite number of operations on the signal $Y(s)$, as shown in [18], [31].

Before performing the backward transform into the time-domain, rearrange the left side terms of (49) using Leibniz' formula for the differentiation of products twice. This results in

$$\frac{1}{s^{N+1}}\frac{d^j}{ds^j}\left(\frac{1}{s}\frac{d^{N-j}}{ds^{N-j}}\left(s^{N+1}Y(s)\right)\right)=\sum_{\kappa_1=0}^{N-j}\sum_{\kappa_2=0}^{j}\binom{N-j}{\kappa_1}\binom{j}{\kappa_2}\times \frac{(N+1)!}{(N-\kappa_1-\kappa_2)!\,(N-\kappa_1+1)}\frac{1}{s^{\kappa_1+\kappa_2+1}}\frac{d^{N-\kappa_1-\kappa_2}}{ds^{N-\kappa_1-\kappa_2}}\,Y(s) \quad (50)$$

which, in view of the right hand side of (49), implies in turn

$$\frac{1}{s^{N+j+2}}\,y^{(j)}(0)=\frac{(-1)^j}{j!\,(N-j)!}\sum_{\kappa_1=0}^{N-j}\sum_{\kappa_2=0}^{j}\binom{N-j}{\kappa_1}\binom{j}{\kappa_2}\times \frac{(N+1)!}{(N-\kappa_1-\kappa_2)!\,(N-\kappa_1+1)}\frac{1}{s^{\kappa_1+\kappa_2+1}}\frac{d^{N-\kappa_1-\kappa_2}}{ds^{N-\kappa_1-\kappa_2}}\,Y(s)\,. \quad (51)$$

Eq. (51) is now transformed back into the time domain. Using the following inverse Laplace transform formulae

$$\pounds^{-1}\left[\frac{1}{s^{i+1}}\frac{d^j}{ds^j}Y(s)\right]=\int_0^t\frac{(t-\tau)^i(-\tau)^j}{i!}\,y(\tau)\,d\tau \quad (52)$$

we obtain

$$\hat{y}^{(j)}(0)=\int_0^t H_j(t,\tau)\,y(\tau)\,d\tau\,,\quad j=0,1,\ldots,N \quad (53)$$

with

$$H_j(t,\tau)=\frac{(N+j+1)!\,(N+1)!\,(-1)^j}{t^{N+j+1}}\times \sum_{\kappa_1=0}^{N-j}\sum_{\kappa_2=0}^{j}\frac{(t-\tau)^{\kappa_1+\kappa_2}\,(-\tau)^{N-\kappa_1-\kappa_2}}{\kappa_1!\kappa_2!(N-j-\kappa_1)!(j-\kappa_2)!(N-\kappa_1-\kappa_2)!(\kappa_1+\kappa_2)!(N-\kappa_1+1)} \quad (54)$$

The results obtained above thus give an estimate $\hat{y}^{(j)}(t)$ at time $t=0$ from the polynomial signal y, see (1), taken on the interval $[0,t]$. In order to get a moving-horizon and causal version of these results, first replace t with $-T$, where T is a positive constant [4], [3] and simplify using the fact that

$$(-1)\;H_j(-T,-\tau)=(-1)^j\;H_j(T,\tau) \quad (55)$$

Finally, by shifting the y-values by t, Theorem 1 is immediate. ■

B. Lemma for the Proof of Theorem 2

Lemma 1 (Inverse of $\mathbf{W}_{\mathrm{r}}(t_0, t_1)$): Let the entries of the matrix $\mathbf{W}_{\mathrm{r}}(t_0, t_1)$ be given as in (24). The entries of its inverse are

$$\left[W_{\mathrm{r}}^{-1}\right]_{ij}(t_0, t_1) = \frac{(i-1)!\,(j-1)!\,(i+j-1)}{(t_1-t_0)^{i+j-1}} \binom{N+i}{N+1-j} \binom{N+j}{N+1-i} \binom{i+j-2}{i-1}^2. \quad (56)$$

□

Proof: In light of equation (24), first, left- and right-multiply $\mathbf{W}_{\mathrm{r}}(t_0, t_1)$ with a diagonal matrix $\mathbf{M}$ whose entries are

$$M_{ij} = \frac{(i-1)!}{(t_0-t_1)^i}\delta_{ij} \quad (57)$$

where δ_{ij} is the Kronecker delta. Then, proceed with computing the following matrix product in component form as

$$\begin{aligned}
&[(t_1-t_0)\,\mathbf{M}\,\mathbf{W}_{\mathrm{r}}(t_0,t_1)\,\mathbf{M}]_{ij} \\
&= (t_1-t_0)\sum_{k=1}^{N+1}\sum_{l=1}^{N+1} M_{ik}\,[W_{\mathrm{r}}]_{kl}\,(t_0,t_1) M_{lj} \\
&= (t_1-t_0)\sum_{k=1}^{N+1}\sum_{l=1}^{N+1} \frac{(i-1)!}{(t_0-t_1)^i}\delta_{ik}\,\frac{-(t_0-t_1)^{k+l-1}}{(k-1)!(l-1)!(k+l-1)}\,\frac{(l-1)!}{(t_0-t_1)^l}\delta_{lj} \\
&= \frac{1}{i+j-1}
\end{aligned} \quad (58)$$

whose result can be recognized as the entries of an $(N+1)\times(N+1)$ Hilbert matrix, hereafter denoted $\mathbf{H}$. The entries of the inverse of $\mathbf{H}$ are known to be [25]

$$\left[H^{-1}\right]_{ij} = (-1)^{i+j}\,(i+j-1)\binom{N+i}{N+1-j}\binom{N+j}{N+1-i}\binom{i+j-2}{i-1}^2 \quad (59)$$

and by computing

$$\mathbf{W}_{\mathrm{r}}^{-1}(t_0,t_1) = (t_1-t_0)\,\mathbf{M}\,\mathbf{H}^{-1}\,\mathbf{M} \quad (60)$$

we obtain (25), which completes the proof of the Lemma. ■

References

[1] R. W. Brockett, *Finite Dimensional Linear Systems*. Wiley, 1970.

[2] C.-T. Chen, *Linear system theory and design*. Oxford Univ. Press, 1999.

[3] M. Fliess, C. Join, M. Mboup, and A. Sedoglavic, "Estimation des dérivées d'un signal multidimensionnel avec applications aux images et aux vidéos," *Actes 20^e Coll. GRETSI*, Louvain-la-Neuve, 2005.

[4] M. Fliess, C. Join, M. Mboup, and H. Sira-Ramírez, “Analyse et représentation de signaux transitoires: application à la compression, au débruitage et à la détection de ruptures,” *Actes 20^e Coll. GRETSI*, Louvain-la-Neuve, 2005.

[5] M. Fliess, C. Join, and H. Sira-Ramírez, “Robust residual generation fo linear fault diagnosis: an algebraic setting with examples,” *Int. Journal of Control*, vol. 77, pp. 1223–1242, 2004.

[6] M. Fliess, C. Join, and H. Sira-Ramírez, “Non-linear estimation is easy,” *Int. Journal of Modelling, Identification and Control*, vol. 3, 2008.

[7] M. Fliess, M. Mboup, H. Mounier, and H. Sira-Ramírez, “Questioning some paradigms of signal processing via concrete examples,” in *Algebraic Methods in Flatness, Signal Processing and State Estimation*, H. Sira-Ramírez, G. Silva-Navarro (Eds.), Innovación Editiorial Lagares, Mexico, pp. 1–21, 2003.

[8] M. Fliess and H. J. Sira-Ramírez, “An algebraic framework for linear identification,” *ESAIM Control Optim. Calc. Variat.*, vol. 9, 2003.

[9] M. Fliess and H. J. Sira-Ramírez, “State reconstructors: a possible alternative to asymptotic observers and Kalman filters,” *Proceedings of CESA*, 2003.

[10] M. Fliess and H. Sira Ramírez, “Control via state estimations of some nonlinear systems,” *IFAC Symposium on Nonlinear Control Systems (NOLCOS 2004)*, Stuttgart, 2004.

[11] R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical System Theory*. McGraw-Hill, 1969.

[12] P. G. Kaminski, A. E. Bryson and S. F. Schmidt, “Discrete square root filtering: a survey of current techniques,” *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 727-736, 1971.

[13] H. Khalil, *Nonlinear systems (2nd ed.)*. Prentice-Hall, New York, 1996.

[14] I. R. Khan and R. Ohba, “New design of full band differentiators based on Taylor series,” *IEE Proceedings Vision, Image & Signal Processing*, vol. 146, pp. 185–189, 1999.

[15] M. Krstić, I. Kanellakopoulos and P. Kokotović, *Nonlinear and adaptive control design*. Wiley Interscience, New York, 1995.

[16] W. H. Kwon, P. S. Kim and P. G. Park, “A receding horizon Kalman FIR filter for linear continuous-time systems,” *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2115–2120, 1995.

[17] M. Mboup, “Parameter estimation via differential algebra and operational calculus,” *in preparation*, 2007.

[18] M. Mboup, C. Join, and M. Fliess, “A revised look at numerical differentiation with an application to nonlinear feedback control,” *15th Mediterranean Conference on Control and Automation (MED’07)*, Athens, Greece, 2007.

[19] A. Neves, M. Mboup, and M. Fliess, “An algebraic identification method for the demodulation of QPSK signal through a convolutive channel,” *European Signal Proc. Conf. (EUSIPCO)*, Austria, Vienna, 2004.

[20] J. O’Reilly, *Observers for Linear Systems*, Academic Press, 1983.

[21] J. Reger, J. Jouffroy, “Algebraic Time-Derivative Estimation and Deadbeat State Reconstruction, ” *Technical Report CGR-07-09*, University of Michigan, USA, http://arxiv.org/abs/0710.0010

[22] F. Riesz and B. Sz.-Nagy, *Functional Analysis*. Fredrick Ungar, New York, 1955.

[23] S. Samadi and A. Nishihara, “Explicit formula for predictive FIR filters and differentiators using Hahn polynomials,” *IEICE Transactions*, vol. E90-A, no. 8, pp. 1511–1518, 2007.

[24] S. Sastry, *Nonlinear systems*. Springer, 1999.

[25] L. R. Savage and E. Lukas, “Tables of inverses of finite segments of the Hilbert matrix ” in *Contributions to the Solutions of Systems of Linear Equations and the Determination of Eigenvalues*, O. Taussky (Editor), National Bureau of Standards Applied Mathematics Series, vol. 39, pp. 105–108, 1954.

[26] E. D. Sontag, *Mathematical Control Theory (2nd ed.)*. Springer, 1998.

[27] C.-C. Tseng, "Digital differentiator design using fractional delay filter and limit computation," *IEEE Transactions on Circuits and Systems–I:Regular Papers*, vol. 52, no. 10, pp. 2248–2259, 2002.

[28] S. Valiviita and O. Vainio, "Delayless differentiation algorithm and its efficient implementation for motion control applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 48, no. 5, pp. 967–971, 1999.

[29] H. Wilf and D. Zeilberger, "An algebraic proof theory for hypergeometric (ordinary and "q") multisum/integral identities," *Inventiones Mathematicae*, vol. 108, pp. 575–633, 1992.

[30] J. C. Willems and S. K. Mitter, "Controllability, observability, pole allocation, and state reconstruction," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 582–595, 1971.

[31] J. Zehetner, J. Reger, and M. Horn, "A Derivative Estimation Toolbox based on Algebraic Methods — Theory and Practice, " *IEEE Int. Conf. on Control Applications*, Singapore, 2007.